Models of sector aircraft counts in the presence of local, regional and airport constraints

Deepak Kulkarni

NASA Ames Research Center

Outline

- **Identification of uncertainties in impact of multiple constraints**
- Multiple constraining factors influencing flows
- Quantile regression approach to identify PDFs
- Evaluation of model
- Case study
- Conclusions

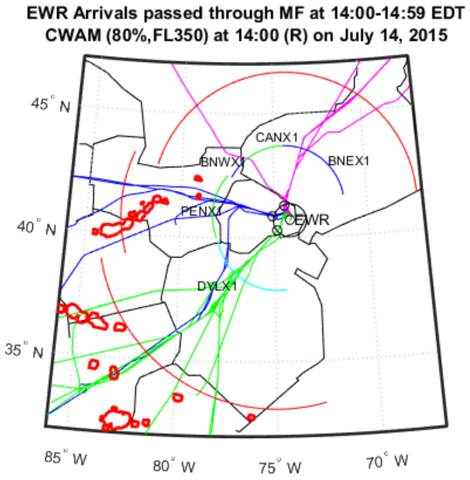
Collaborative Trajectory Option Programs (CTOP)

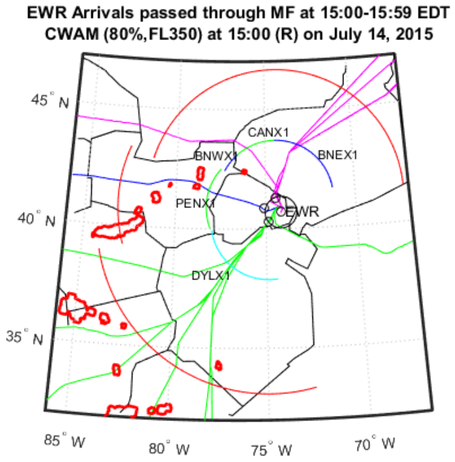
- CTOP is a new traffic management initiative for controlling traffic through ground delays and rerouting

Traffic managers can create multiple flow constrained areas (FCAs)

Flight operators to express and exercise preferred routing options

With Trajectory Option Sets (TOS)

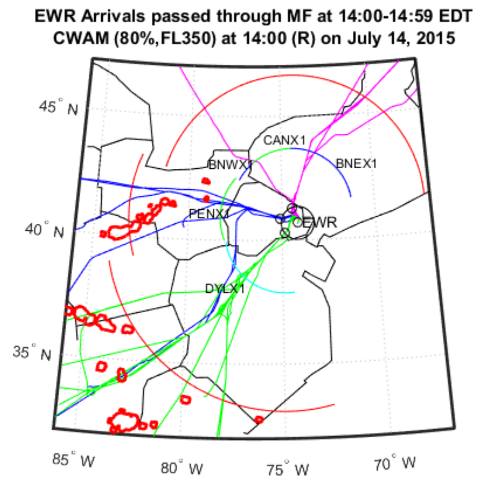




Multiple flow constraints

CTOP Decision Tasks for FAA controllers

- Lack of availability of decision support tools is one of the obstacles in implementing CTOP.
- Tasks involved in running a CTOP involve identifying areas of demand-capacity imbalance, setting and revising rates for FCAs and considering alternative options
- Single biggest challenge in doing this is to reason in the presence of uncertainty.
- Models characterizing maximal flows and **probability distribution functions** of flows and counts in sectors and FCAs in the presence of **multiple constraints** would be useful in creating decision support tools

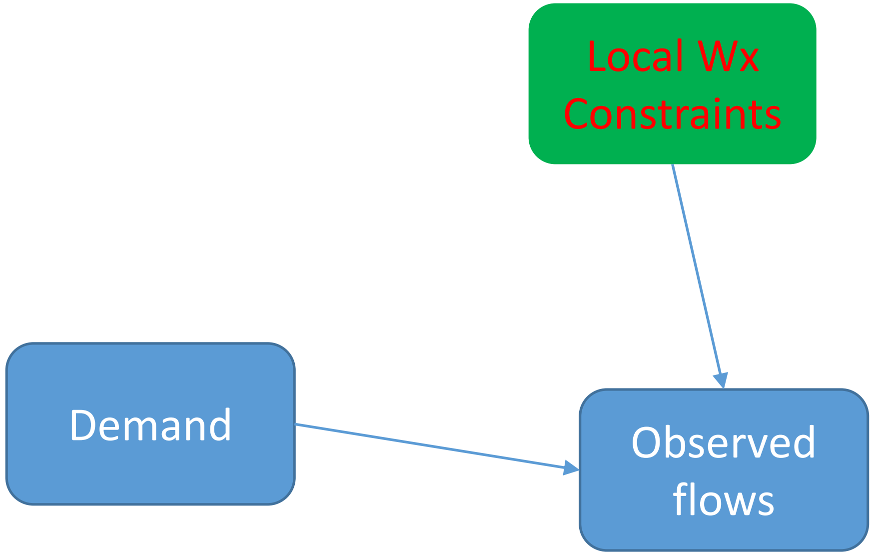


What is the optimal location for FCAs/ rates?

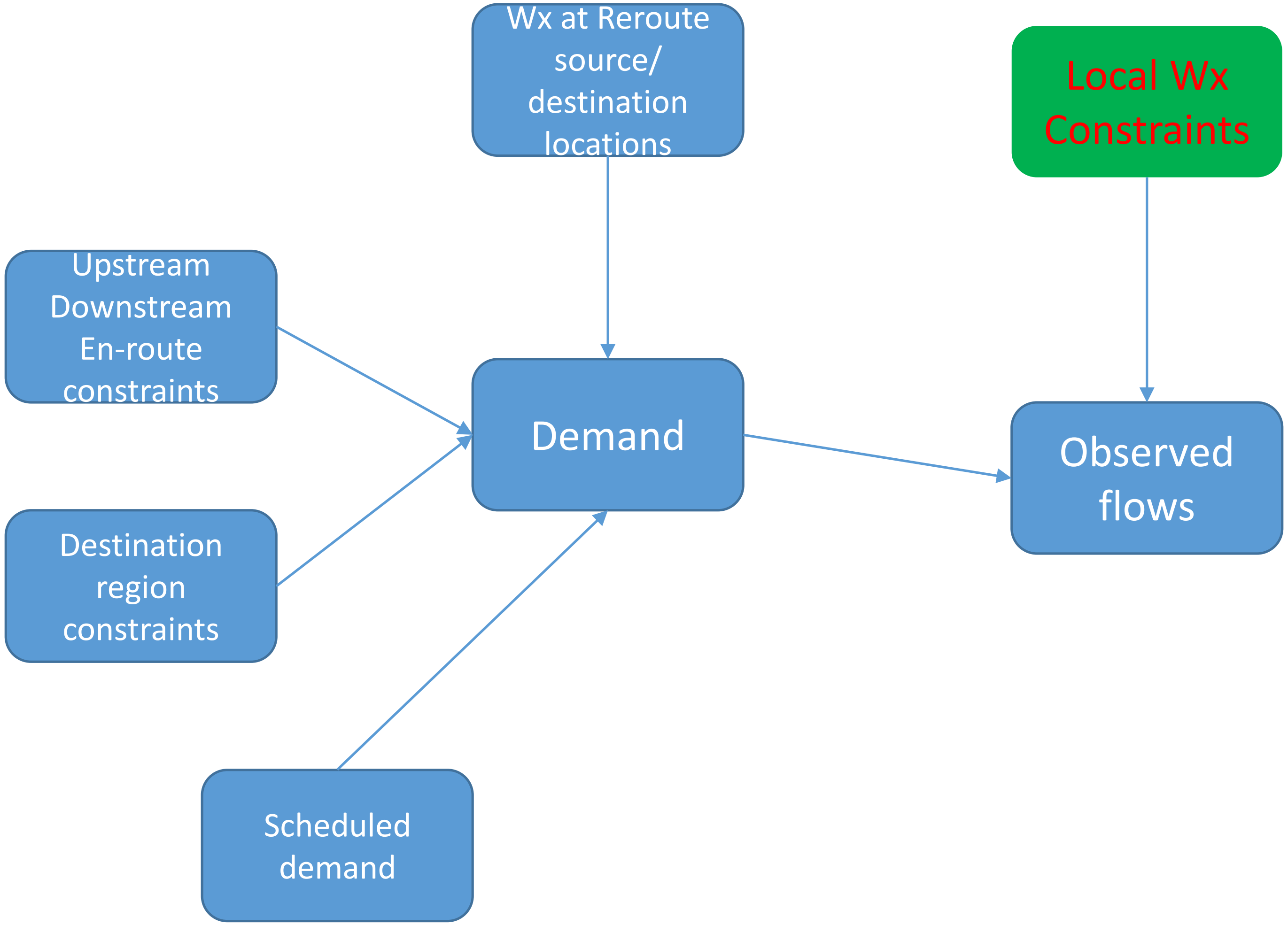
Outline

- Identification of uncertainties in impact of multiple constraints
- **Multiple constraining factors influencing flows**
- Quantile regression approach to identify PDFs
- Evaluation of model
- Case study
- Conclusions

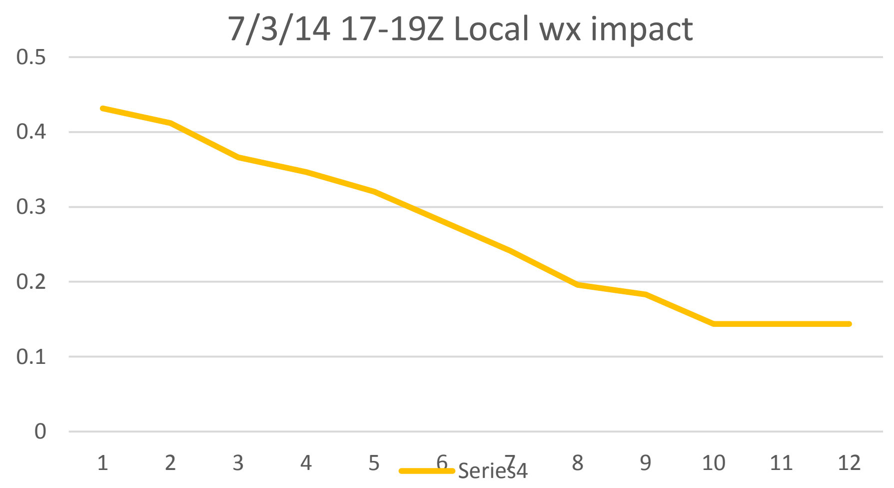
Spatially distributed factors impacting local flows in a sector

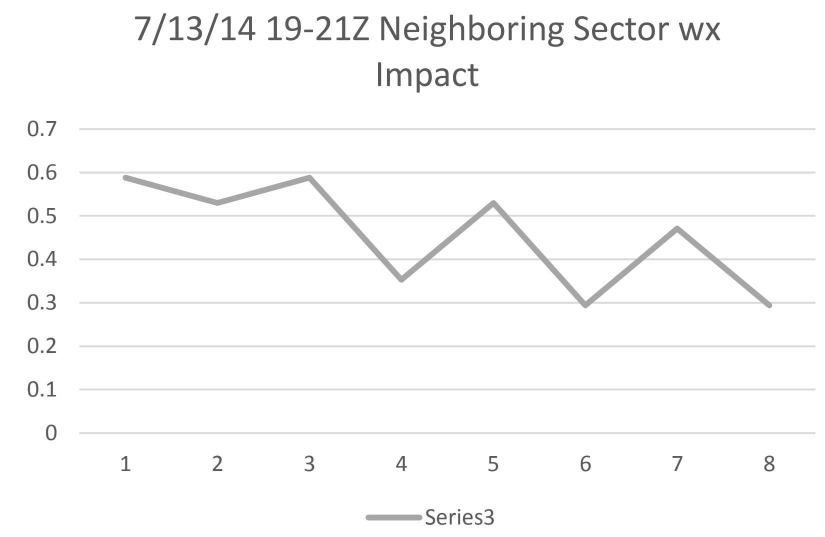


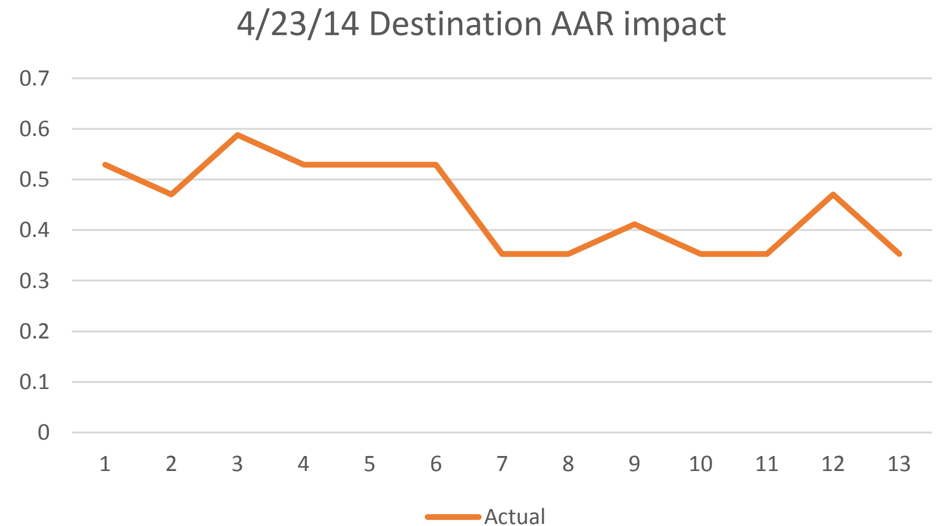
Spatially distributed factors impacting local flows in a sector

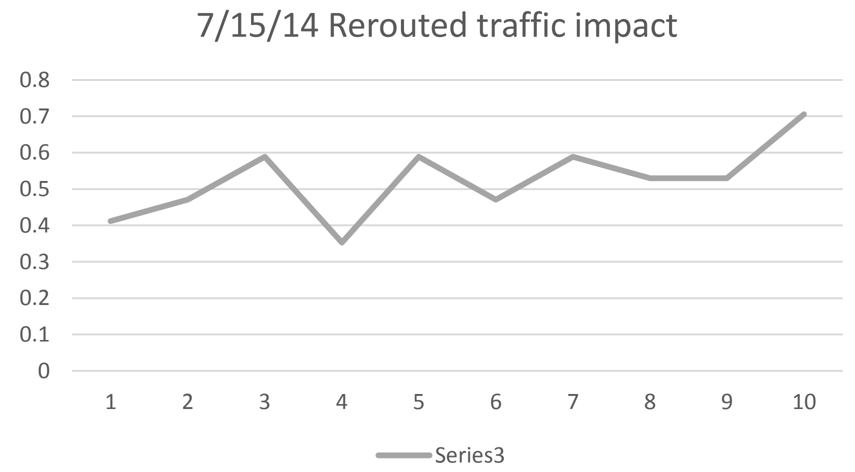


ZNY75 aircraft counts impacted by various factors







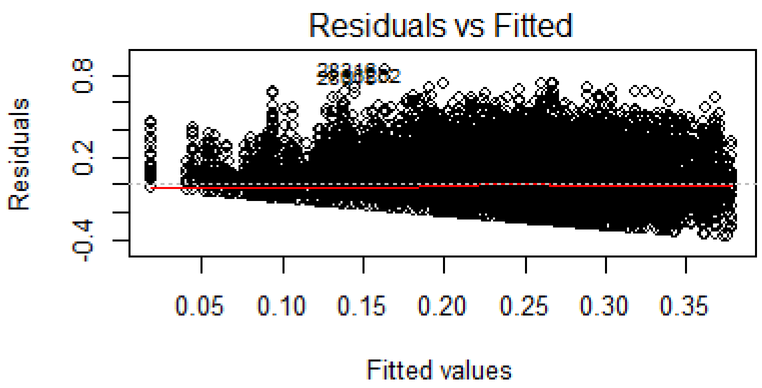


Outline

- Identification of uncertainties in impact of multiple constraints
- Multiple constraining factors influencing flows
- **Quantile regression approach to identify PDFs**
- Evaluation of model
- Case study
- Conclusions

Identification of probability distribution functions

- Accurate characterization of probability distribution function would be useful
- Under Beush-Pagan test, p-value is less than .05 indicating heteroscedasticity for base flow model in terms of weather
- Graphical plot of residual vs fitted values also do not show a random distribution.
- Constraining weather factor would be expected to have impact at higher percentiles and not at lower percentiles.
- Thus approaches such as linear regression won't be suited for this problem

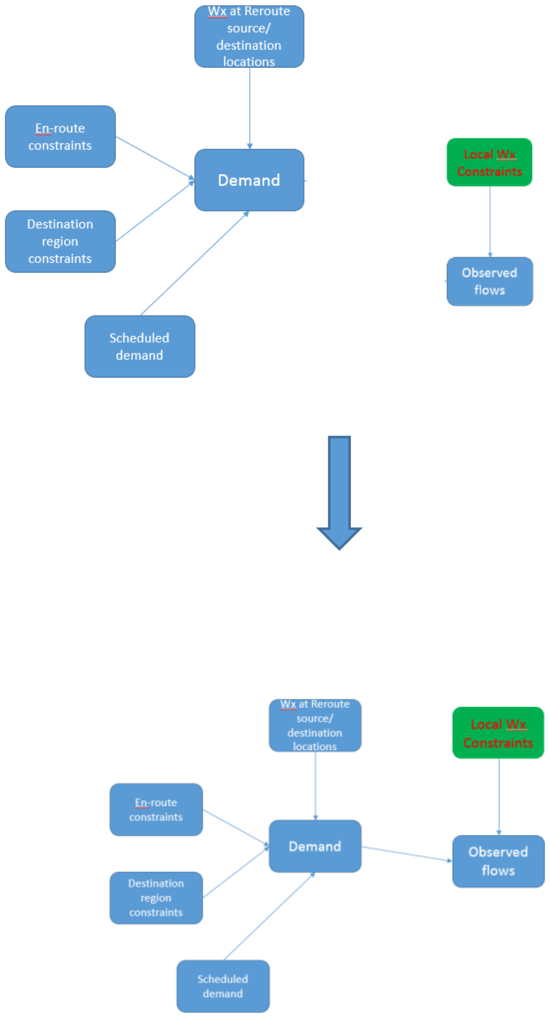


Quantile regression

- Model of n-th percentile of dependent variable.
- Advantages
  - More comprehensive analysis of relationship between dependent variable and independent variables e. g. 95^{th} percentile of observed flows may depend on local weather, but mean values may not be affected by weather when weather is limited. Traditional methods would not capture this.

Models of factors impacting flows

- Lack of sufficient data in high weather conditions under multi-factor conditions is a challenge. So, an approach is taken to decompose the problem.
- Maximum demand model: baseline maximum flow models (g_{dem}) in the presence of clear local weather. This is a function of enroute region constraints, airport constraints and weather in locations that can be source or destination of re-routes to the location of interest
- Weather constrained flow reduction model($f_{wx\text{-}red}$): a model of reduction in baseline counts as a function of local weather.
- The composite model that combines these two can be represented as $f_{wx\text{-}red}$ (g_{dem}(**e**,**a**), **l**) where **e** represents external weather, **a** represents airport constraints and **l** represents local weather.



Models of demand under multiple constraints

Sector demand model

- Data used is quarter hours from the period April-September 2014. Accuracy of statistical method is dependent on amount of data available. To increase the amount of data available, a generalized model is created for sectors in ZNY and ZOB.
- Dependent variable is observed sector counts scaled relative maximum observed sector count.
- Independent variables used are
  - Destination airport AAR (NYCAAR)
  - Weather Impact index for enroute regions (WITIn)
  - Weather Impact Index for regions that are source/ destination of re-routes (WITIs, WITId)

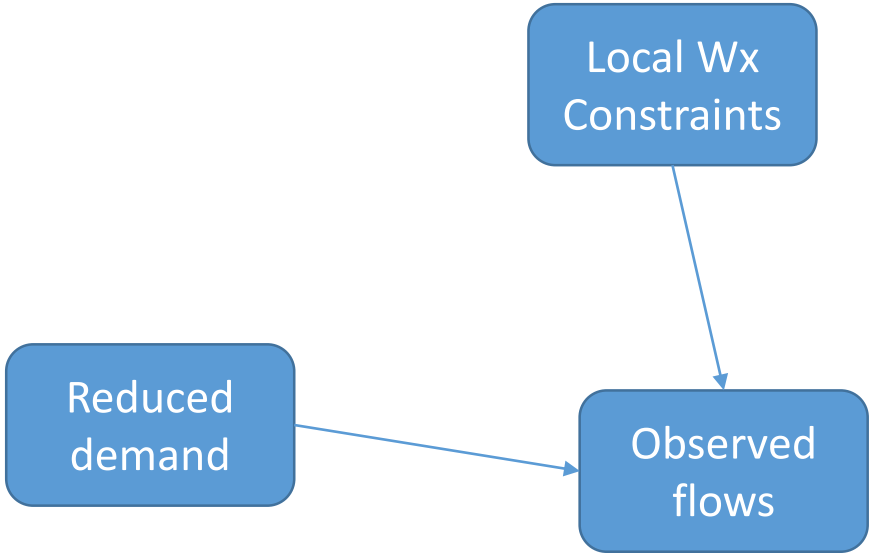
Quantile regression model of sector demand

- Data used: April-September 2014 17-22Z (peak period)
- Dependent variable: Scaled sector counts
- Independent variables used are
  - Destination airport AAR (AAR)
  - Weather Impact index for enroute regions (WITIn)
  - Weather Impact Index for regions that are source/ destination of re-routes (WITIs, WITId)
- All variables are statistically significant p-values in a model of 95th percentile as can be seen in the table below.
- Variable dependencies changes depending which percentile is characterized. 20th percentile is dependent on airport AARs but not other variables.

		Std Error	T value	Pr (> |t|)
WITIn	-.0012	.0004	-2.9	.004
AAR	.0051	.0000	183	.000
WITIs	.0027	.0005	5.3	.000
WITId	-.0007	.0003	-2.5	.014

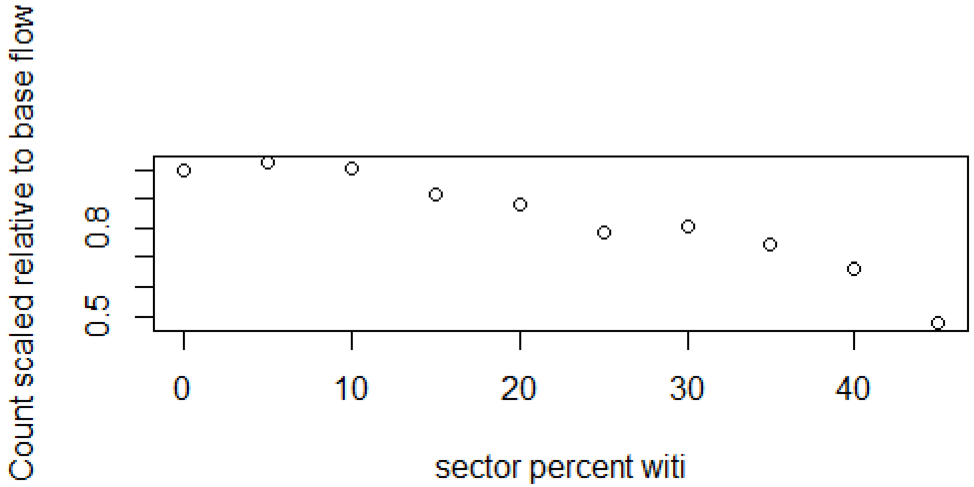
Models of impact of local weather

Impact of multiple constraints on maximum observed flows



Impact of local sector weather

- Sector percent witi bins of width 5 are created starting with [0,5) going to [45,50)
- Scaled sector counts are relative to demand
- X axis show percent witi in a sector.
- Y axis shows 95th percentile value among the scaled sector counts when sector percent witi is in the range shown on the x axis.
- Three different types of models: Point estimation model, Empirical model, theoretical model



Linear model

- Linear regression model is
  - 95th percentile scaled counts are: **1.07 - .011 * percent witi**
  - R-squared = .9
- A theoretical model (reduction in capacity is equal to percent witi) is **1 - .01 * percent witi**
- These model differ only slightly

```
Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)        1.073956   0.033761  31.811 1.04e-09 ***
cutoffs1_10[1:10] -0.010906   0.001265  -8.623 2.54e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.05744 on 8 degrees of freedom
Multiple R-squared:  0.9029, Adjusted R-squared:  0.8907 
F-statistic: 74.35 on 1 and 8 DF,  p-value: 2.536e-05
```

Model errors

- A test data set is used to calculate errors in different approaches to estimate 95th values.
- On test data, linear model has average error of .03 and theoretical model has an average error of .07
- Error in point estimates varies from 0 for clear weather to .21 for heavy weather..

Lower bound of witi	Upper bound of witi	Number of points	95th percentile in test data	95th percentile of observed count	Lower bound of quantile estimate	Upper bound of quantile estimate	Predicted count with linear model	Predicted count with theoretical model
0	5	75733	1.00	1.00	1.00	1.00	1.07	1.00
5	10	1905	1.01	1.02	0.99	1.09	1.02	0.95
10	15	1029	1.02	1.00	0.98	1.09	0.96	0.90
15	20	459	0.88	0.92	0.86	1.13	0.91	0.85
20	25	321	1.01	0.88	0.78	1.07	0.86	0.80
25	30	204	0.88	0.78	0.68	0.99	0.80	0.75
30	35	141	0.82	0.80	0.65	0.99	0.75	0.70
35	40	54	0.69	0.75	0.54	0.97	0.69	0.65
40	45	51	0.71	0.66	0.59	0.87	0.64	0.60

Comparison of different types of models

- Theoretical model has worse error compared to empirical model.
- Theoretical model can be used more broadly in all situations – FCAs and sectors not studied and rare weather situations
- Point estimation models has least amount of errors in low weather conditions where there is a lot of data.
- Not enough data in heavy weather to distinguish between models that differ mainly in heavy weather situations

Composite model

- Dependent variable: Scaled sector counts
- Independent variables used are
  - Destination airport AAR (AAR)
  - Weather Impact index for enroute regions (WITIn)
  - Weather Impact Index for regions that are source/ destination of re-routes (WITIs, WITId)
  - Local weather (witi_local)
- The composed model :
  - witi_local * (a*AAR + b* WITIn + c*WITIs + d*WITId)

Outline

- Identification of uncertainties in impact of multiple constraints
- Multiple constraining factors influencing flows
- Quantile regression approach to identify PDFs
- **Evaluation of model**
- Case study
- Conclusions

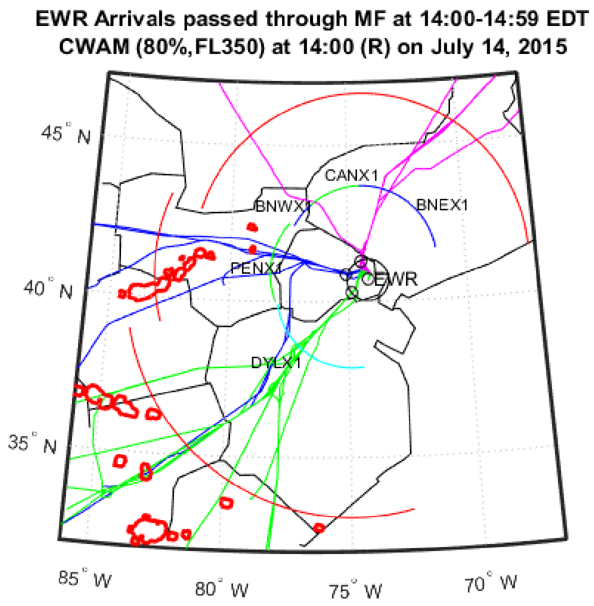
Evaluation of Composite model

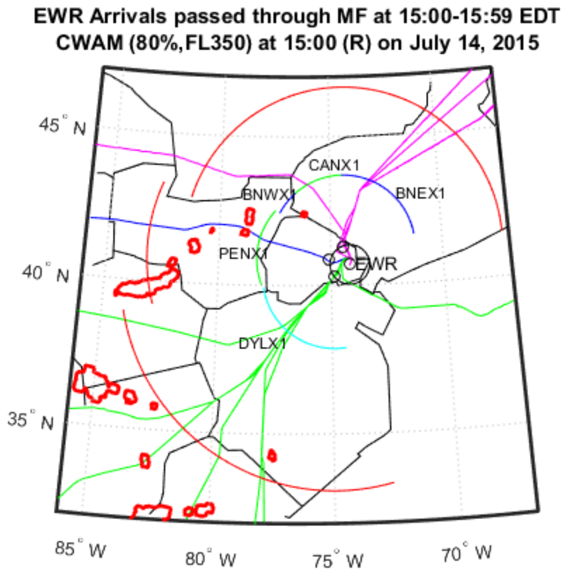
- On a test data sample, it would be expected that about 95% of data would fall below values predicted by this model and about 5% of the time model would under-predict observed counts.
- Three month data was used for testing this model.
- On this data, 92% of observed counts were below the model prediction and 8% of counts were above the model prediction.

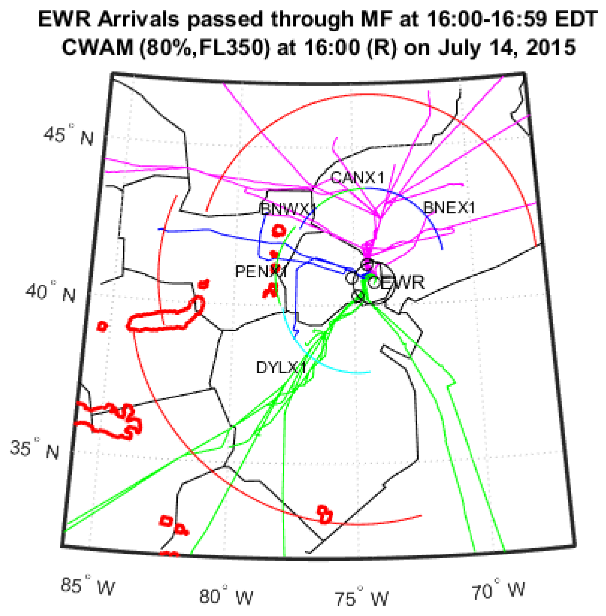
Outline

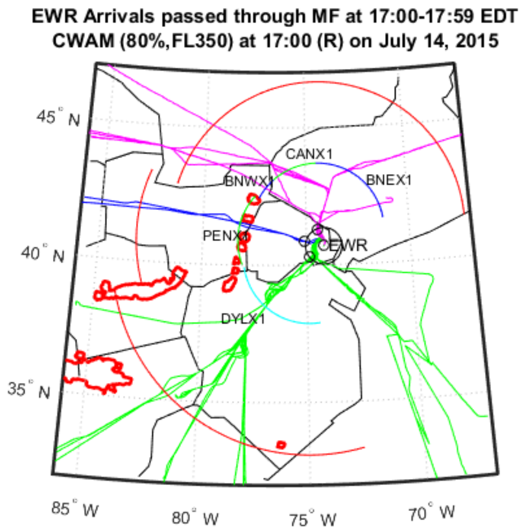
- Identification of uncertainties in impact of multiple constraints
- Multiple constraining factors influencing flows
- Quantile regression approach to identify PDFs
- Evaluation of model
- **Case study**
- Conclusions

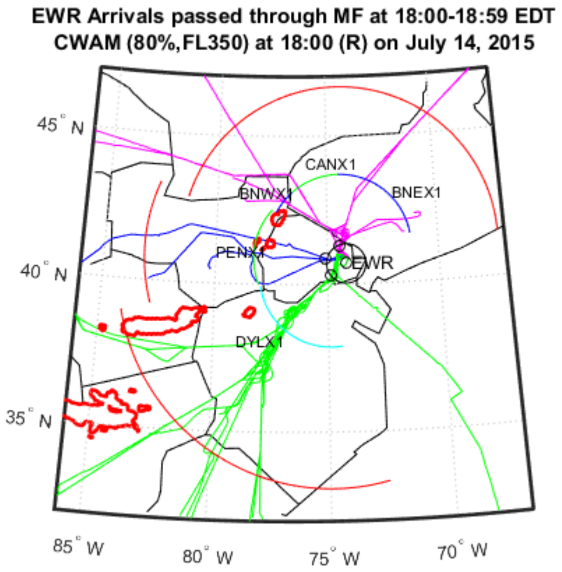
Case study day: July 14, 2015









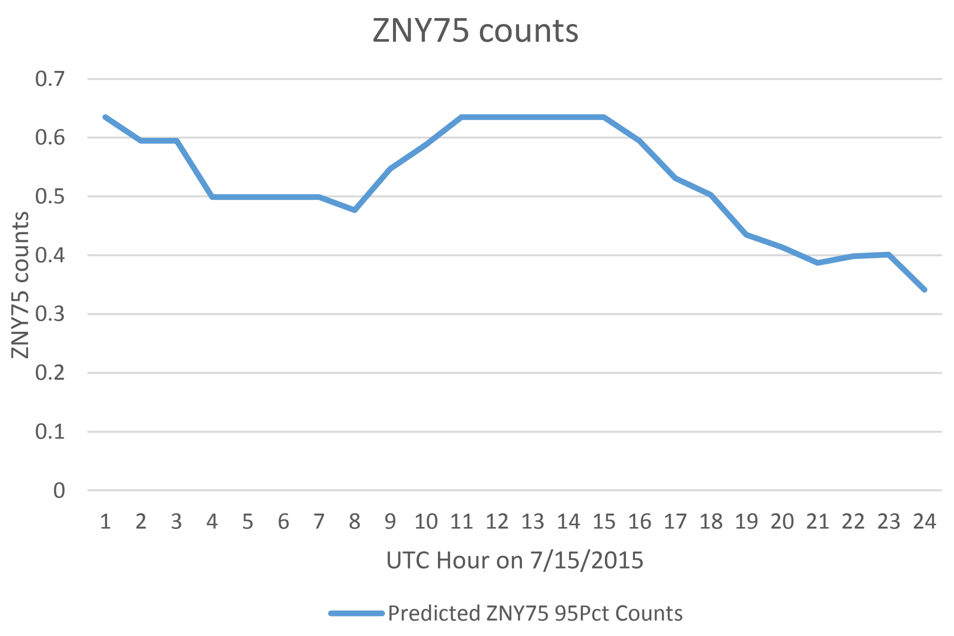


NOWCAST weather with EWR traffic

ZNY75

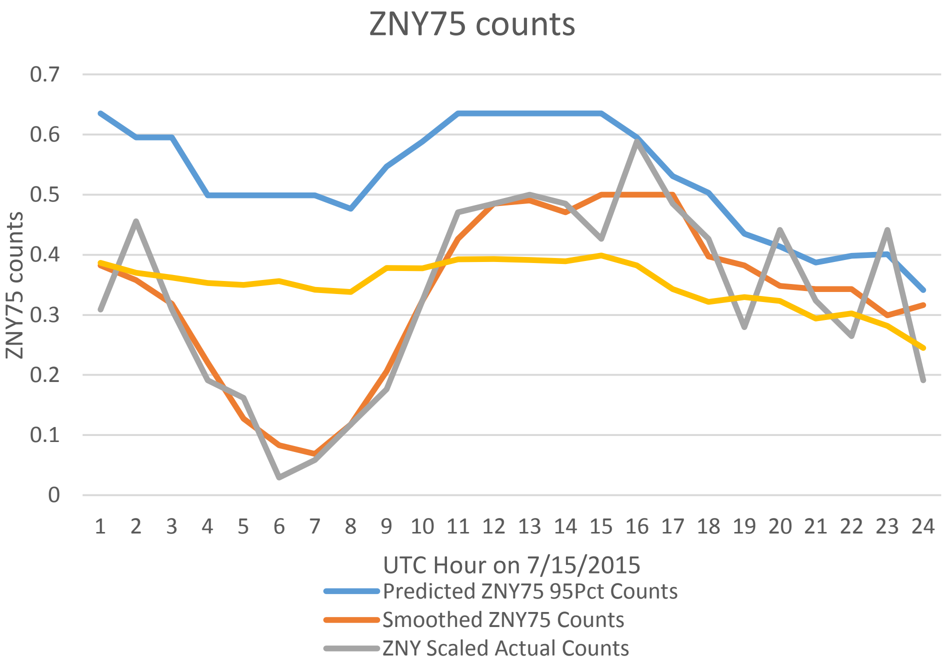
- Expected allowed flows drop by 37%
- Comparison with demand probability distribution indicates very high probability of demand-capacity imbalance needing FCA.

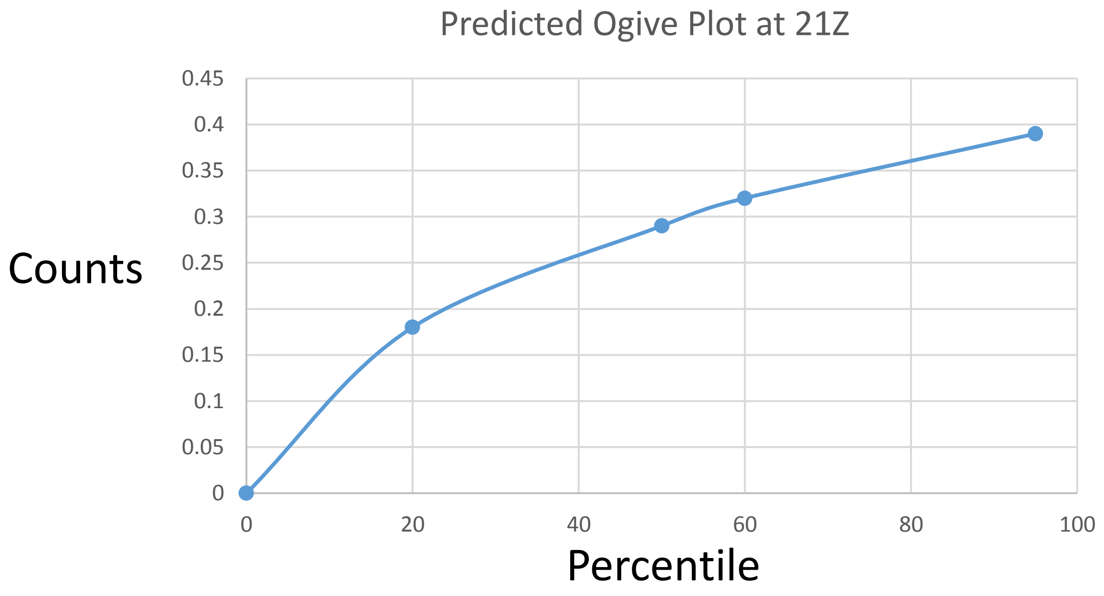
Scaled ZNY75 counts



Flows on 6-21-16 2000

Ogive plot





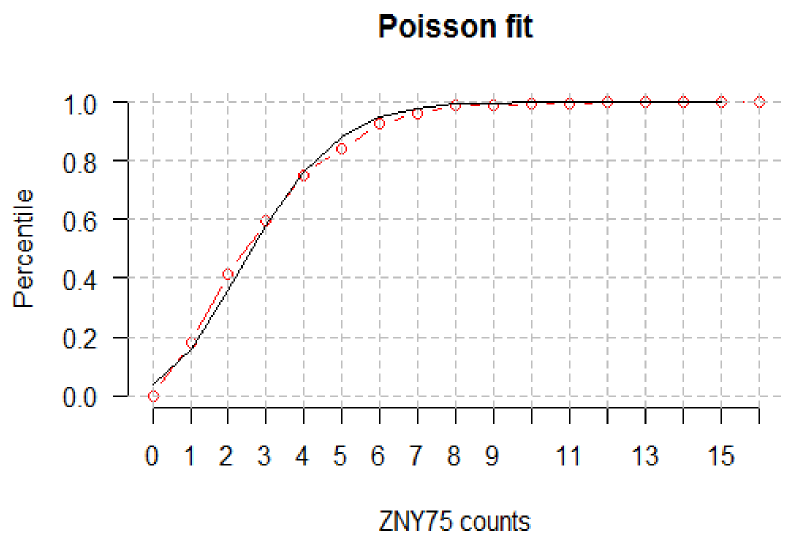
7/14/15 21Z counts at 60th percentile
4/15/14 19Z counts at 95th percentile

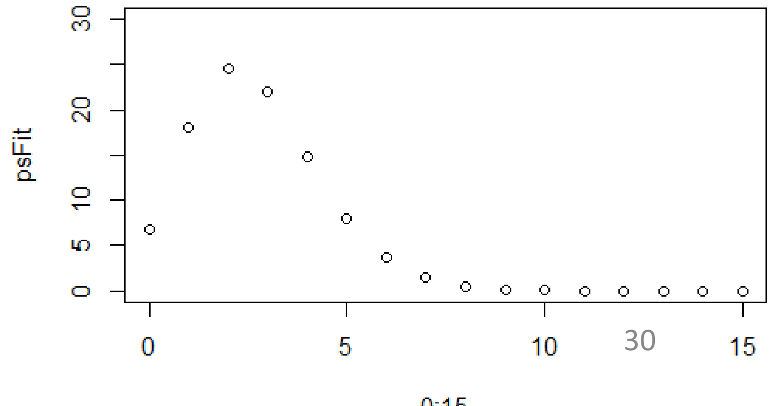
Fit to Poisson distribution

Probability distribution function can be used with appropriate stochastic optimization algorithms to identify optimal FCA rates

Closed form probability distribution to fit a series of percentiles

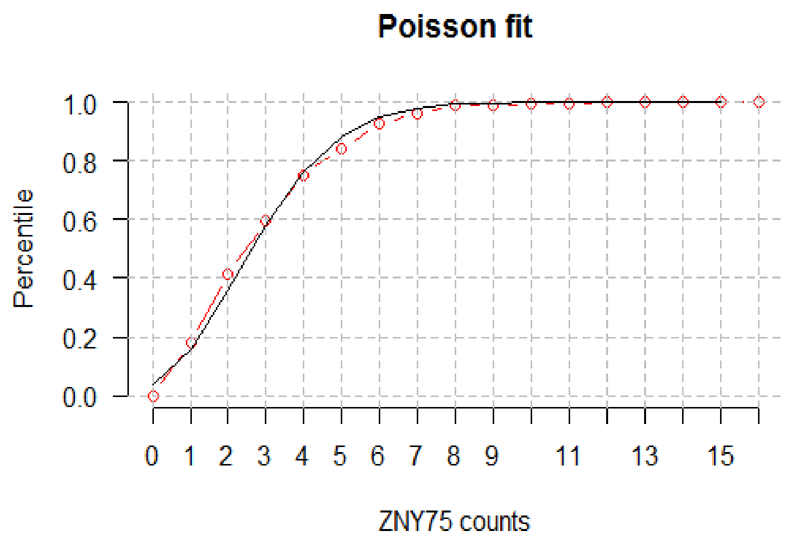
- Appropriate distribution function fitting quantile estimates can be created
- Figure on the top shows a Poisson cumulative distribution function (lambda = 2.8) fit to a series of quantile estimates of ZNY75 counts with ZNY local WITI at .25
- Note that Peak hour distribution is closer to normal.

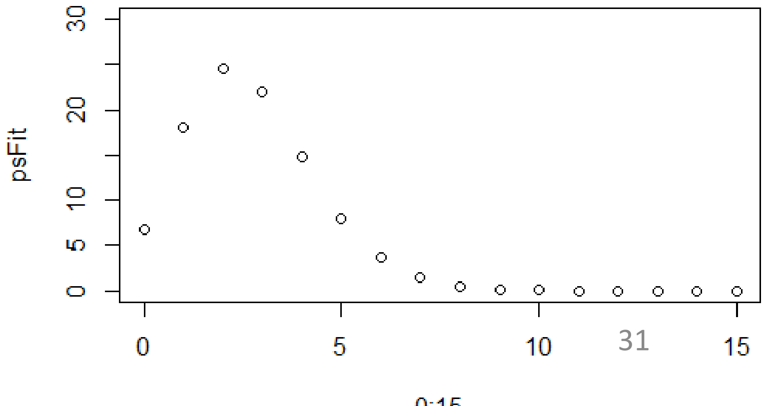




Potential uses of demand probability distribution functions

- If actual weather is worst than forecast weather used to set CTOP parameters, there will be demand overage. On the other hand, If actual weather is better than forecast weather used to set CTOP parameters, there will be aircraft that were unnecessarily subjected to delays or re-routing.
- In the context of what-if-reasoning, different parameters can be derived. Relevant derived parameters can be computed. For example, parameter of interest can be expected value (constrained-flow – F) when demand > F.
- In the example shown, If F is set to be 5, demand overage is .22 which is 10% of aircraft instantaneous count.
- Closed form demand distribution can also be used in optimization to estimate optimal flow rates





Outline

- Identification of uncertainties in impact of multiple constraints
- Multiple constraining factors influencing flows
- Quantile regression approach to identify PDFs
- Evaluation of model
- Case study
- **Conclusions**

Conclusion

- Identification of uncertainties in impact of multiple constraints would be useful in creating CTOP DSTs.
- Following factors have statistically significant influence: Destination airport constraints, local weather, neighboring region weather, weather at reroute source/destinations
- Quantile models are useful in accurate characterization of probability distribution functions.
- Decomposed model approach was taken to create accurate models in the presence of limited data.